wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 7, 2025 – 10:33 AM EDT

PDB ID : 9E1J / pdb_00009e1j
Title : Alpha-Delta heterodimeric form of soluble hydrogenase I from Pyrococcus furiosus. Data processed and model refined in P21221
Authors : Lanzilotta, W.N.; Adams, M.W.W.
Deposited on : 2024-10-21
Resolution : 2.60 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

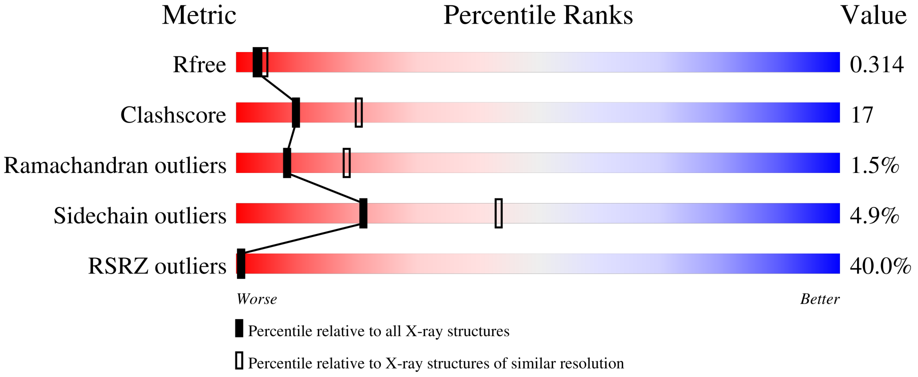
MolProbity : 4-5-2 with Phenix2.0rc1
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0rc1
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.006 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.44

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

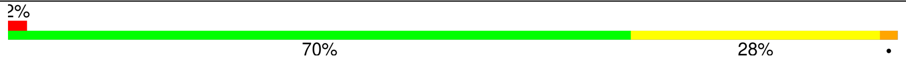
The reported resolution of this entry is 2.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	3775 (2.60-2.60)
Clashscore	180529	4181 (2.60-2.60)
Ramachandran outliers	177936	4129 (2.60-2.60)
Sidechain outliers	177891	4129 (2.60-2.60)
RSRZ outliers	164620	3775 (2.60-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	256	
1	C	256	
2	B	419	
2	D	419	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	SF4	A	501	-	-	X	-
3	SF4	C	502	-	-	X	-
4	FCO	B	501	-	-	X	-

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 10787 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Sulfhydrogenase 1 subunit delta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	256	Total 2009	C 1292	N 328	O 369	S 20	0	1	0
1	C	254	Total 1987	C 1276	N 325	O 366	S 20	5	0	0

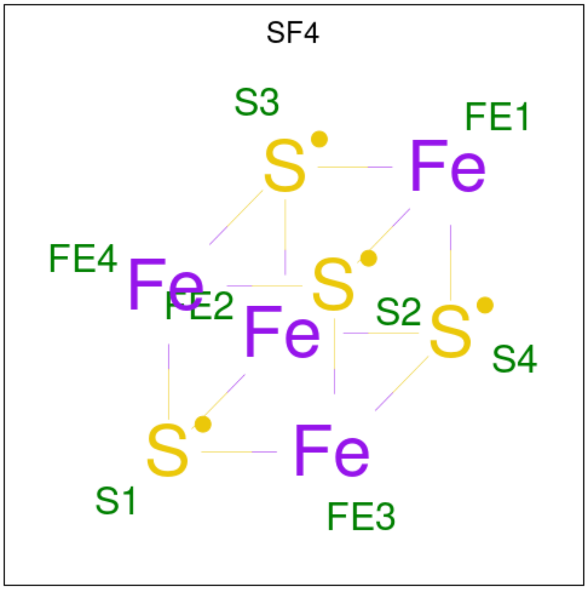
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	226	ALA	LYS	conflict	UNP E7FHU4
A	252	ALA	GLU	conflict	UNP E7FHU4
C	226	ALA	LYS	conflict	UNP E7FHU4
C	252	ALA	GLU	conflict	UNP E7FHU4

- Molecule 2 is a protein called Sulfhydrogenase 1 subunit alpha.

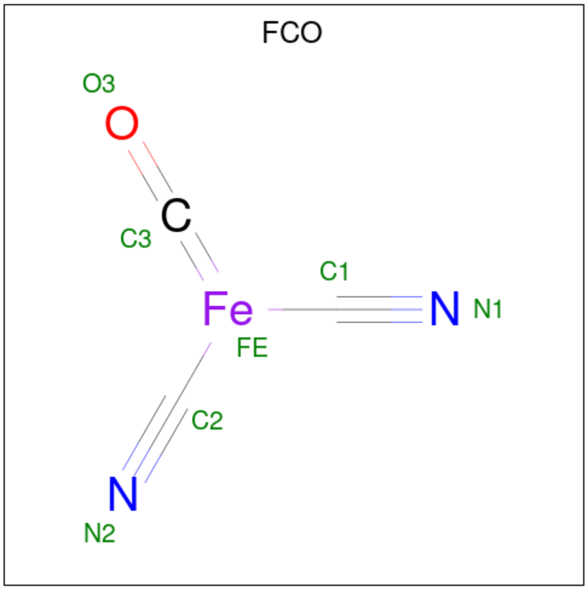
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	419	Total 3332	C 2145	N 556	O 616	S 15	0	1	0
2	D	419	Total 3324	C 2140	N 554	O 615	S 15	0	0	0

- Molecule 3 is IRON/SULFUR CLUSTER (CCD ID: SF4) (formula: Fe_4S_4).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Fe S 8 4 4	0	0
3	A	1	Total Fe S 8 4 4	0	0
3	A	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0
3	C	1	Total Fe S 8 4 4	0	0

- Molecule 4 is CARBONMONOXIDE-(DICYANO) IRON (CCD ID: FCO) (formula: C_3FeN_2O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	Fe	N	O		
4	B	1	7	3	1	2	1	0	0

- Molecule 5 is NICKEL (II) ION (CCD ID: NI) (formula: Ni).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ni		
5	B	1	1	1	0	0
5	D	1	1	1	0	0

- Molecule 6 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Cl		
6	B	1	1	1	0	0

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
7	A	31	31	31	0	0
7	B	30	30	30	0	0

3 Residue-property plots ⓘ

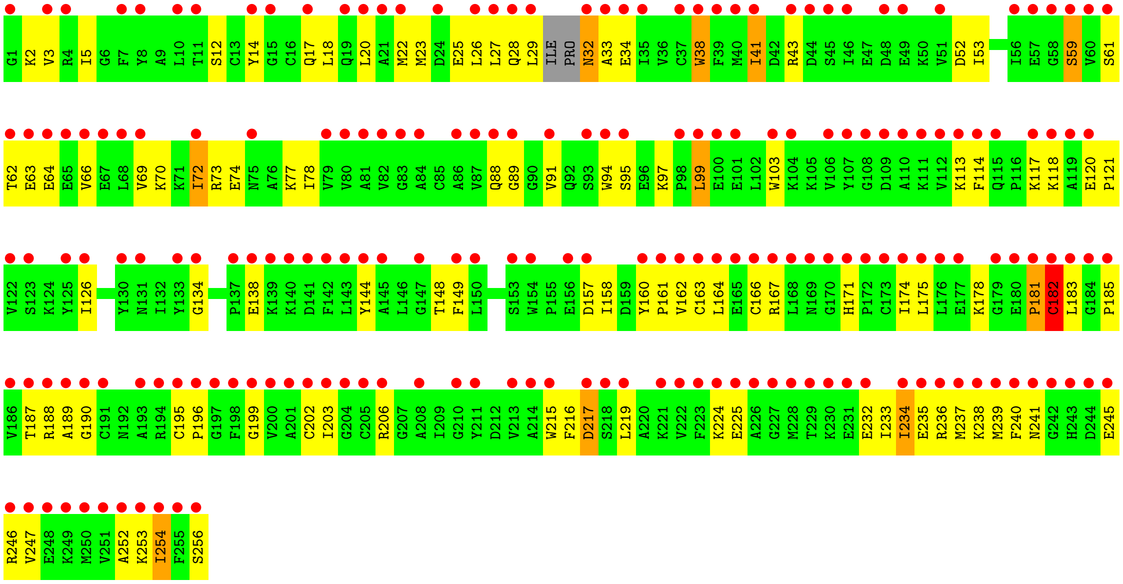
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

- Molecule 1: Sulfhydrogenase 1 subunit delta

Chain A:

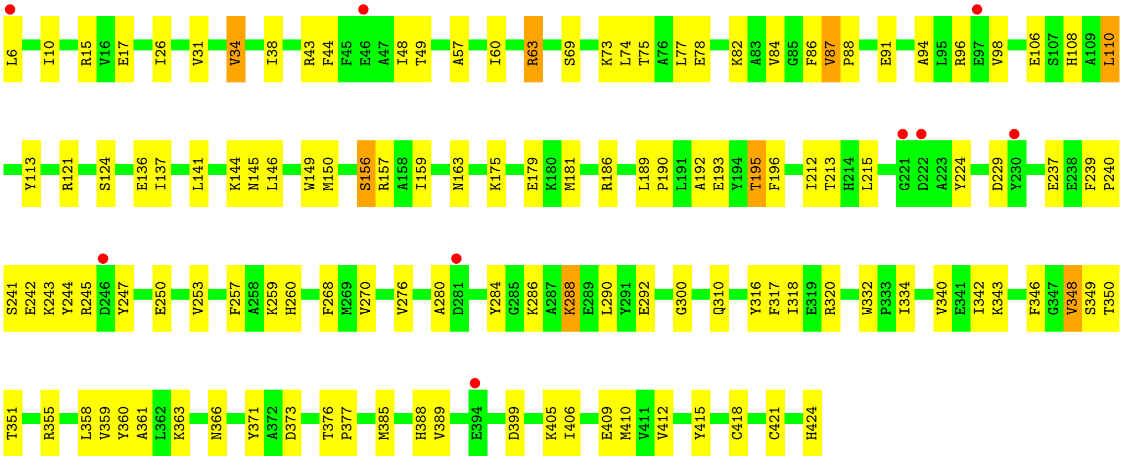
- Molecule 1: Sulfhydrogenase 1 subunit delta

Chain C:



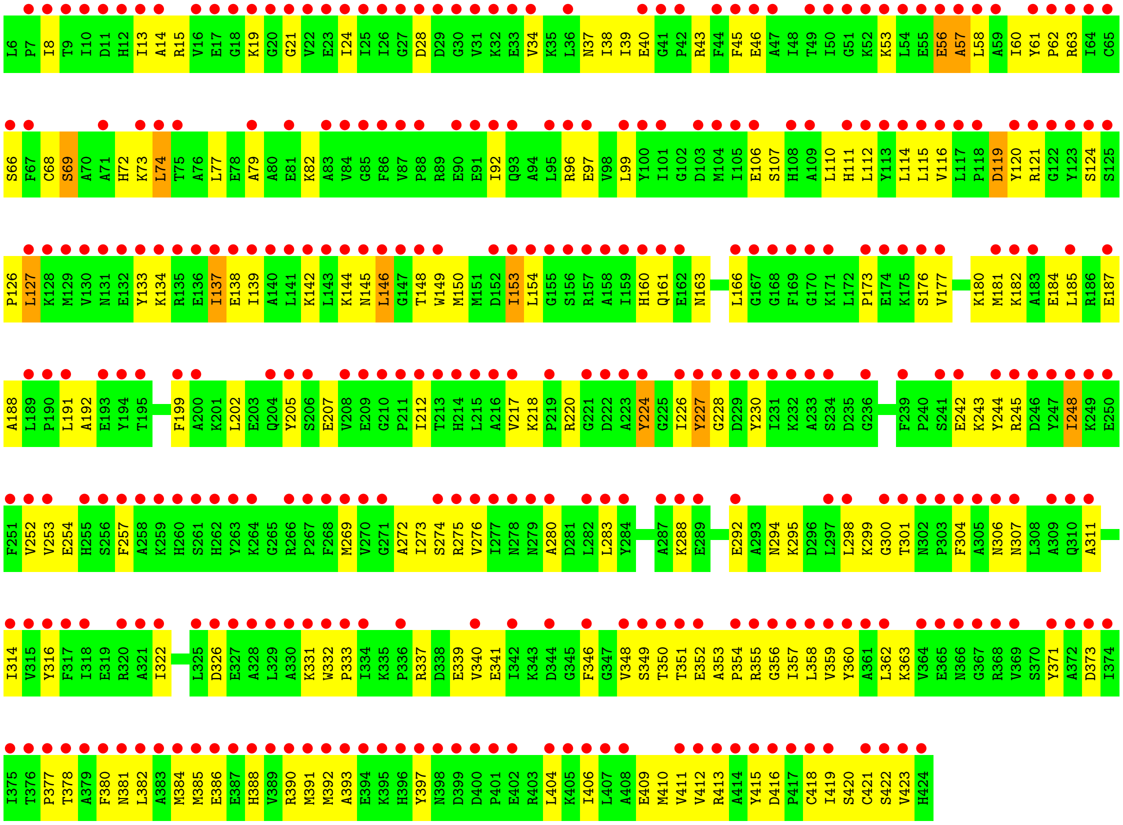
- Molecule 2: Sulfhydrogenase 1 subunit alpha

Chain B:



- Molecule 2: Sulfhydrogenase 1 subunit alpha

Chain D:



4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 2 21	Depositor
Cell constants a, b, c, α, β, γ	94.29Å 111.19Å 141.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.91 – 2.60 35.91 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (35.91-2.60) 93.3 (35.91-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.86 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.256 , 0.313 0.258 , 0.314	Depositor DCC
R_{free} test set	44333 reflections (4.33%)	wwPDB-VP
Wilson B-factor (Å^2)	34.5	Xtriage
Anisotropy	0.835	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 66.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.88	EDS
Total number of atoms	10787	wwPDB-VP
Average B, all atoms (Å^2)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 23.67 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 4.4544e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: SF4, CL, FCO, NI

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| >5$	RMSZ	$\#\|Z\| >5$
1	A	0.53	0/2055	0.73	0/2769
1	C	0.39	0/2028	0.68	0/2732
2	B	0.57	0/3406	0.77	1/4599 (0.0%)
2	D	0.34	0/3394	0.57	0/4584
All	All	0.47	0/10883	0.69	1/14684 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	63	ARG	CA-C-O	5.08	127.04	118.91

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts ⓘ

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2009	0	2012	59	0
1	C	1987	0	1980	82	0
2	B	3332	0	3359	93	0
2	D	3324	0	3352	156	0
3	A	24	0	0	4	0
3	C	40	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	7	0	0	8	0
5	B	1	0	0	0	0
5	D	1	0	0	0	0
6	B	1	0	0	0	0
7	A	31	0	0	2	0
7	B	30	0	0	4	0
All	All	10787	0	10703	369	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 17.

The worst 5 of 369 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:239:MET:HG3	2:B:60:ILE:HD11	1.43	1.00
2:B:377:PRO:HD2	4:B:501:FCO:N1	1.80	0.96
1:C:43:ARG:HH12	2:D:406:ILE:HG22	1.31	0.96
2:B:377:PRO:HD2	4:B:501:FCO:C1	1.97	0.94
2:D:19:LYS:HB3	2:D:40:GLU:HG2	1.54	0.89

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	255/256 (100%)	245 (96%)	10 (4%)	0	100	100
1	C	250/256 (98%)	214 (86%)	27 (11%)	9 (4%)	3	4
2	B	418/419 (100%)	392 (94%)	21 (5%)	5 (1%)	11	24
2	D	417/419 (100%)	387 (93%)	24 (6%)	6 (1%)	9	19
All	All	1340/1350 (99%)	1238 (92%)	82 (6%)	20 (2%)	8	18

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	181	PRO
1	C	182	CYS
1	C	234	ILE
2	D	56	GLU
2	D	57	ALA

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	216/215 (100%)	205 (95%)	11 (5%)	20	42
1	C	213/215 (99%)	198 (93%)	15 (7%)	12	27
2	B	351/350 (100%)	341 (97%)	10 (3%)	38	65
2	D	350/350 (100%)	331 (95%)	19 (5%)	18	39
All	All	1130/1130 (100%)	1075 (95%)	55 (5%)	21	43

5 of 55 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	157	ASP
1	C	254	ILE
2	D	421	CYS
2	D	248	ILE
1	C	182	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	28	GLN
2	D	160	HIS
2	D	366	ASN
2	D	279	ASN
2	B	366	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 12 ligands modelled in this entry, 3 are monoatomic - leaving 9 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	FCO	B	501	2	0,6,6	-	-	-		
3	SF4	C	504	-	0,12,12	-	-	-		
3	SF4	C	502	1	0,12,12	-	-	-		
3	SF4	A	501	1	0,12,12	-	-	-		
3	SF4	C	505	-	0,12,12	-	-	-		
3	SF4	A	503	1	0,12,12	-	-	-		
3	SF4	A	502	1	0,12,12	-	-	-		
3	SF4	C	503	1	0,12,12	-	-	-		
3	SF4	C	501	1	0,12,12	-	-	-		

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SF4	C	504	-	-	-	0/6/5/5

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SF4	C	502	1	-	-	0/6/5/5
3	SF4	A	501	1	-	-	0/6/5/5
3	SF4	C	505	-	-	-	0/6/5/5
3	SF4	A	503	1	-	-	0/6/5/5
3	SF4	A	502	1	-	-	0/6/5/5
3	SF4	C	503	1	-	-	0/6/5/5
3	SF4	C	501	1	-	-	0/6/5/5

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	501	FCO	8	0
3	C	502	SF4	2	0
3	A	501	SF4	2	0
3	A	503	SF4	1	0
3	A	502	SF4	1	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	256/256 (100%)	0.19	2 (0%)	82	79	19, 41, 60, 75	1 (0%)
1	C	253/256 (98%)	3.77	198 (78%)	0	0	6, 36, 51, 69	250 (98%)
2	B	419/419 (100%)	0.18	9 (2%)	63	58	22, 36, 56, 83	1 (0%)
2	D	419/419 (100%)	3.45	330 (78%)	0	0	8, 38, 54, 70	409 (97%)
All	All	1347/1350 (99%)	1.87	539 (40%)	1	1	6, 37, 56, 83	661 (49%)

The worst 5 of 539 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	13	ILE	13.3
1	C	255	PHE	12.8
2	D	264	LYS	12.7
1	C	226	ALA	12.0
2	D	51	GLY	11.8

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no oligosaccharides in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	SF4	C	504	8/8	0.57	0.20	107,117,129,141	8
3	SF4	C	505	8/8	0.63	0.20	54,78,100,103	8
3	SF4	C	502	8/8	0.77	0.17	49,64,76,84	8
3	SF4	C	503	8/8	0.86	0.11	59,74,85,88	8
3	SF4	C	501	8/8	0.90	0.12	49,58,73,75	8
6	CL	B	503	1/1	0.90	0.11	66,66,66,66	0
5	NI	B	502	1/1	0.92	0.06	52,52,52,52	0
4	FCO	B	501	7/7	0.95	0.18	8,9,11,13	0
5	NI	D	501	1/1	0.97	0.06	61,61,61,61	1
3	SF4	A	501	8/8	0.98	0.05	22,23,25,26	0
3	SF4	A	502	8/8	0.98	0.04	20,24,27,27	0
3	SF4	A	503	8/8	0.98	0.04	22,27,29,32	0

6.5 Other polymers ⓘ

There are no such residues in this entry.